

The logic of the self-refutation argument in *Dissoi Logoi* 4.6

Sebastiano Molinelli

Dissoi Logoi 4.6 presents a beautiful self-refutation argument, which I analyse here, offering a different assessment of its relation to self-contradiction and the Liar paradox from the only one available in the literature.

The text to be examined is as follows:

αἱ γάρ τις ἐρωτάσαι τὼς λέγοντας ὡς ὁ αὐτὸς λόγος εἴη ψεύστας καὶ ἀλαθῆς ὃν αὐτοὶ λέγοντι, πότερος ἐστίν· αἱ μὲν “ψεύστας”, δᾶλον ὅτι δύο εἴη· αἱ δ’ “ἀλαθῆς” ἀποκρίναιτο, καὶ ψεύστας ὁ αὐτὸς οὗτος. (*Dissoi Logoi* 4.6)¹

Although this is one of the oldest, if not the oldest, testimonies of self-refutation argument in Western philosophy² and has often been compared with the Liar paradox,³ only Luca Castagnoli has devoted adequate space to its logic in his monograph on self-refutation in 2010, concluding that the argument leverages self-refutation, but not self-contradiction, and hence that it has mere dialectical value, but not a logical one.⁴ In this

¹ This is the text established by Thomas M. Robinson in his authoritative edition (Robinson, T. M. (ed.) (1979), *Contrasting arguments. An Edition of the Dissoi Logoi*, New York: Arno Press, 124) and which, in this passage, has been confirmed in Laks and Most's recent one (Laks, A. and Most, G. W. (eds.) (2016), *Early Greek Philosophy, Volume IX: Sophists, Part 2*, Cambridge (MA): LOEB, 188-190).

² Out of the many questions that the work carries, one concerns its composition date, the common belief being that ‘the Δ. Λ. was written some time around 403-395’ (Robinson 1979, 41). But cf. Mazzarino, S. (1966), *Il pensiero Storico Classico*. Vol. 1, Bari: Laterza, 151 and Maso, S. (ed.) (2018), *Dissoi Logoi. Edizione criticamente rivista, Introduzione, Traduzione, Commento*, Roma: Edizioni di Storia e Letteratura, 38-39 for an earlier dating, whereas Freeman, K. (ed.) (1946), *The Pre-Socratic philosophers: a companion to Diels, 'Fragmente der Vorsokratiker'*, Oxford: Basil Blackwell, 417, n. a1 and Conley, T. M. (1985), ‘Dating the So-Called *Dissoi Logoi*: A Cautionary Note’, *Ancient Philosophy* 5: 59-65, at 62-63 for a later one.

³ Robinson 1979, 193-194, Desbordes, F. (1987), ‘Aux origines de la linguistique: l'exemple des *Dissoi logoi*’, in S. Mellet (ed.), *Etudes de linguistique générale et de linguistique latine offertes en hommage à Guy Serbat*, Paris: Peeters Publishers, 33-43, at 39, Maso 2018, 120-121.

⁴ Castagnoli, L. (2010), *Ancient Self-Refutation. The Logic and History of the Self-Refutation Argument from Democritus to Augustine*, Cambridge: Cambridge University Press, at 28-29. He defines a self-refutation argument as ‘any argument which aims at showing that (and how) something is “self-refuting”, i.e. refutes itself’ (ibid., 3), while he takes ‘self-contradiction to include all those cases in which a single proposition, atomic or compound, either entails or consists of a pair of contradictory propositions’ (ibid., 5). Although he acknowledges that the

paper, I will challenge this view, also on the basis of a logical flaw that I have spotted in Castagnoli's reconstruction.

To begin with, in order to make things clearer, I will adopt Castagnoli's own translation of the passage, which runs thus:

For if one were to ask those who say that the same λόγος is false and true which of the two their own λόγος is, if <their reply were> 'false', it is clear that <the false and the true λόγος> would be two things, while if they were to answer 'true', then this very <λόγος> would be also false.⁵

Likewise, I will borrow his label 'Identity Thesis' (abbr. 'IT') for

'the same λόγος is false and true',

and I will add just a new one, 'Difference Thesis' ('DT'), for

'the false and the true λόγος would be two things'.⁶

After having defended IT in the first part of the chapter (§§ 4.2-5), now the anonymous sophist author of the *Dissoi Logoi* is arguing for DT (§§ 4.6-9), and his first argument is the one under examination here.

The author starts by imagining that the IT supporter is asked whether his own λόγος is true or false. On the first horn of this dilemma, the IT supporter answers that his own λόγος is false, and this leads the author to conclude DT. But a logical difficulty then rises, because, on the one hand, through his answer the IT supporter certainly concedes 'the contradictory of his IT (as long as he endorses the platitude $Fp \rightarrow \neg p$)',⁷ to quote Castagnoli.⁸ On the other hand, *contra* Castagnoli,⁹ the contradictory of IT is not DT, DT being the *contrary* of IT, instead. I shall show how this is so by moving from the

edges between these two notions 'are not always as sharp as we might desire' (ibid., 7), he stresses how self-contradictions are 'necessary falsehoods [...] and are rejected as such in most logical systems' (ibid., 5-6), whereas self-refutation arguments do not 'prove, or aim to prove, the falsehood of the thesis which incurs defeat' (ibid., 355).

⁵ Ibid., 25.

⁶ In other words, the false and the true λόγος are two numerically distinct (δύο) objects, or better, sets of objects, as we will see soon (see *infra*, n. 11).

⁷ With '*p*' being a λόγος, and '*Fp*' standing for the predicate '*p* is false'.

⁸ Ibid., 27-28.

⁹ '[...] i.e. that, to borrow the opaque but now familiar jargon of the *Dissoi Logoi*, the false λόγος and the true λόγος are two different things' (ibid., 28).

formalization proposed by Castagnoli himself.¹⁰ He correctly paraphrases IT as ‘any λόγος whatsoever is (unqualifiedly) both false and true’,¹¹ which he formalizes as

$$(\forall p) (Tp \wedge Fp),$$

with ‘ p ’ being a λόγος, ‘ Tp ’ standing for the predicate ‘ p is true’, and ‘ Fp ’ for ‘ p is false’. If that is the case, then $F(IT)$, namely ‘it is not true that any λόγος whatsoever is (unqualifiedly) both false and true’, would be of the form

$$\neg (\forall p) (Tp \wedge Fp),$$

which is equivalent to ‘there is at least one λόγος that is not (unqualifiedly) both true and false’, namely

$$(\exists p) \neg (Tp \wedge Fp),$$

or, by the negation of the conjunction rule, to ‘there is at least one λόγος that is (unqualifiedly) either non-true or non-false’

$$(\exists p) (\neg Tp \vee \neg Fp).$$

However, this being an inclusive disjunction, it is true also in the case in which both $\neg Tp$ and $\neg Fp$ are true, whereas ‘nothing in our text suggests that the author of the *Dissoi Logoi* envisaged the possibility of truth-value gaps’,¹² just as he never alludes to the possibility of intermediate states when dealing with the other couples of opposite attributes discussed in the work, namely good/bad (chapter 1), beautiful/ugly (ch. 2), just/unjust (ch. 3), insane/sane and wise/ignorant (ch. 5). It is hence necessary, first, to turn this inclusive disjunction into the exclusive one

¹⁰ Ibid., 25-26.

¹¹ Throughout the chapter, starting from § 4.1, the author frequently uses individual terms such as ‘the same λόγος’, ‘the false λόγος’, ‘the true λόγος’ not as definite descriptions — as the definite article would at first sight suggest — but as universally quantified, and hence as equivalent to their corresponding universal terms, such as ‘any λόγος whatsoever’/‘all λόγοι’, ‘any false λόγος whatsoever’/‘all false λόγοι’, ‘any true λόγος whatsoever’/‘all true λόγοι’. This is the case also here, and both IT and DT are hence to be taken as universal propositions (cf. Stebbing, L. S. (1930), *A Modern Introduction to Logic*, London: Methuen, 79, 149, where the analogous example ‘The whale is a mammal’ is given).

¹² Ibid., 26, n. 14, where Castagnoli adds that this is also the reason why he has preferred $(\forall p) (Tp \wedge Fp)$ over $(\forall p) (Tp \leftrightarrow Fp)$ as formalization of IT.

$(\exists p) (\neg Tp \vee \neg Fp)$,

then, by the same token, to replace $\neg Tp$ with Fp , and $\neg Fp$ with Tp , so obtaining:

$(\exists p) (Tp \vee Fp)$.

This reads ‘there is at least one λόγος that is (unqualifiedly) exactly either true or false’, and I will call it ‘PDT’ from now on, where ‘P’ alludes to its nature of particular predication. PDT is clearly different from DT, which expresses a universal proposition¹³ and must therefore be paraphrased, first, as ‘no λόγος whatsoever is (unqualifiedly) both false and true’, then — again bearing in mind that truth-value gaps seem to be excluded — as ‘no λόγος whatsoever is (unqualifiedly) both false and true, nor neither of these’, or, more compactly, as ‘any λόγος whatsoever is (unqualifiedly) exactly either false or true’. Its formalization is

$(\forall p) (Tp \vee Fp)$.

Notwithstanding the big difference between PDT and DT, the two being subalterns, the author is not afraid of straining logic and concluding the latter unduly, instead of the former legitimately, from the negation of IT. After all, what a sophist like him seeks in this dilemma is first of all support for DT, no matter if merely rhetorical at this stage.

Proceeding with the second horn, we encounter the self-refutation proper. The IT supporter answers that his own λόγος is true, and the author points out the paradoxical consequence that this answer entails, namely that IT, being a λόγος itself, must be false too. In formal terms, ‘if they were to answer “true”’ can be expressed as

$T(IT)$.

Next, the author understands two steps. First, if IT is true, then IT is the case (semantic descent):

$T(IT) \rightarrow (IT)$.

Then, since IT is a λόγος, by self-application of IT and, hence, substitution of the variable p in $(\forall p) (Tp \wedge Fp)$, we obtain

¹³ See *supra*, n. 11.

$T(IT) \wedge F(IT)$.

This conjunction is what the author expressly concludes through ‘then this very <λόγος> would be also false’, where ‘also’ (καί in Greek) implicitly indicates that $T(IT)$ too, although left understood, comes with $F(IT)$ as the outcome of this second branch of the reasoning. Furthermore, such a conclusion represents the simplest and clearest case of contradiction, boiling down to the form ‘ $p \wedge \neg p$ ’. But if so, then the author has proved that the assumption of $T(IT)$ entails a contradiction, which is tantamount to saying that he has refuted IT by *reductio ad contradictionem*. Therefore, although not openly stating so, the conclusion at which he arrives is $F(IT)$, and the path to get it can be abridged as

$T(IT) \rightarrow F(IT)$.

As a result, this whole dilemma is not to be viewed, as Castagnoli argues, simply as a ‘dialectical silencer’¹⁴ of IT , with the aim of pointing out the ‘dialectical defeats’¹⁵ that the thesis inevitably encounters ‘as soon as it is posed under scrutiny’.¹⁶ On the contrary, it should be noted how the refutation of IT in the second horn is also logically effective, as it shows how the IT supporter is bound to violate the principle of non-contradiction. It is not by chance that precisely against the deniers of this principle Aristotle himself will use, among others, a proof that is very similar to our argument, but so far neglected in the literature on *Dissoi Logoi*, namely:

But if everyone equally both is in error and states the truth, there will be nothing for such a person to speak or say; for he simultaneously says this and not this.¹⁷ (Arist. *Metaph.* Γ 4, 1008b7-10)

In sum, *contra* Castagnoli, who excludes that our argument means ‘to prove the necessary falsehood’¹⁸ of IT and who keeps self-refutation and self-contradiction separate,¹⁹ the

¹⁴ Ibid., 29.

¹⁵ Ibid., 28.

¹⁶ Ibid., 29.

¹⁷ Translation in Kirwan, C. (tr. and comm.) (1993), *Aristotle's Metaphysics Books Gamma, Delta, and Epsilon*, Oxford: Clarendon Press, 14 from the Greek εἰ δὲ ὁμοίως ἅπαντες καὶ ψεύδονται καὶ ἀληθῆ λέγουσιν, οὔτε φθέγγασθαι οὔτ' εἰπεῖν τῷ τοιούτῳ ἔσται: ἅμα γὰρ ταῦτά τε καὶ οὐ ταῦτα λέγει.

¹⁸ Castagnoli 2010., 28.

¹⁹ See *supra*, n. 4.

author of the *Dissoi Logoi* shows that IT refutes itself exactly by reason of its being a self-contradiction, and, hence, a ‘necessary falsehood[s]’, to quote Castagnoli himself.²⁰

Furthermore, as the scholar highlights, self-refutation arguments must be assessed also in consideration of their rhetorical aims.²¹ From this perspective, it is then possible to spot a single plan underlying our dilemma and to indicate a way to reconcile it. *First of all*, we must recall that the whole second part of the chapter (§§ 4.6-9) is devoted to supporting DT. *Secondly*, it is reasonable to think that precisely in order to conclude, although invalidly, in favour of DT the author sets up the first horn of the dilemma: for this, if taken in itself, would otherwise be odd, moving from, and not towards, the falsity of IT, contrary to what one would expect a credible attack on a thesis to do. As a result of these two premises, it is reasonable to think, *contra* Castagnoli, that at the height of the second horn of the dilemma the author is highly interested in demonstrating the falsity of IT,²² because that has just been proved to be a way to get T(DT) too. But if that is the case, then it would not be hazardous to think that the tacit F(IT), with which the second horn concludes, is the key for a last, additional, and — again — understood logical step, by which to connect the two parts of the argument, so far kept apart. In formal terms, we would, in fact, have:

- | | |
|-------------------------------|-------------------------------------|
| (1) $F(IT) \rightarrow T(DT)$ | First horn (invalid); |
| (2) $T(IT) \rightarrow F(IT)$ | Second horn; |
| (3) $T(IT) \rightarrow T(DT)$ | From (2) and (1), by concatenation. |

Granted, this reconstruction is speculative and does not autonomously emerge from the text. Nonetheless, by showing how the truth of IT entails not only its self-refutation, but also the truth of the rival DT, this reading would justify the presence, unique in the work, of such a dilemma with the goal of §§ 4.6-9 themselves, namely to argue for DT.

Finally, this argument has been compared with the Liar paradox, a long debated one, whose ancient origins go back to Eubulides of Miletus (D.L. 2.108), whose oldest

²⁰ Castagnoli 2010, 5.

²¹ Ibid., 16 et passim.

²² Cf. ibid., 28. Castagnoli argues that ancient self-refutation arguments in general ‘did not aim at establishing the truth value of a certain proposition’ (ibid., 15-16).

testimony seems to be at Cic. Ac. 2.95,²³ and whose many formulations can be reduced to the form

‘I am speaking falsely’.²⁴

Castagnoli’s denial of the similarity between the two arguments on the grounds that *Dissoi Logoi*’s one is not equally conceived to prove the truth value of the thesis at stake²⁵ should be revised, because such an intention does seem to belong to our author too, as we have just seen. The difference may be found, rather, in the fact that at the end of our argument the sentence in question receives a precise truth value, unlike the Liar, which is a paradox precisely because it does not. For, on the one hand, IT proves to be nothing but false, because F(IT) follows from T(IT) itself, but the converse is not the case. On the other hand, the truth value of the Liar cannot be decided on the basis of the principle of bivalence, because if it is assumed to be false, then it turns out to be true, and *vice versa*. This asymmetry is therefore crucial to draw a line between our dilemma and the Liar, as the latter is characterised exactly by double truth value reversal, whereas our author is so far from concluding T(IT) by force of F(IT), that he, rather, chooses to go in the contrary direction, invalidly deriving T(DT).

Sebastiano Molinelli
Independent Researcher

References

- Castagnoli, L. (2010), *Ancient Self-Refutation. The Logic and History of the Self-Refutation Argument from Democritus to Augustine*, Cambridge: Cambridge University Press.
- Conley, T. M. (1985), ‘Dating the So-Called *Dissoi Logoi*: A Cautionary Note’, *Ancient Philosophy* 5: 59-65. [DOI: 10.5840/ancientphil19855128]

²³ See Mignucci, M. (1999), ‘Paradoxes’, in K. Algra, J. Barnes, J. Mansfeld, and M. Schofield (eds.), *The Cambridge History of Hellenistic Philosophy*, Cambridge: Cambridge University Press, 157-176, 163.

²⁴ See also Castagnoli 2010, 81, n. 46, 82, n. 47.

²⁵ Ibid., 15-16, 28-29 (esp. n. 19).

Desbordes, F. (1987), 'Aux origines de la linguistique: l'exemple des *Dissoi logoi*', in S. Mellet (ed.), *Etudes de linguistique générale et de linguistique latine offertes en hommage à Guy Serbat*, Paris: Peeters Publishers, 33-43.

Freeman, K. (ed.) (1946), *The Pre-Socratic philosophers: a companion to Diels, 'Fragmente der Vorsokratiker'*, Oxford: Basil Blackwell.

Kirwan, C. (tr. and comm.) (1993), *Aristotle's Metaphysics Books Gamma, Delta, and Epsilon*, Oxford: Clarendon Press.

Laks, A. and Most, G. W. (eds.) (2016), *Early Greek Philosophy, Volume IX: Sophists, Part 2*, Cambridge (MA): LOEB.

Maso, S. (ed.) (2018), *Dissoi Logoi. Edizione criticamente rivista, Introduzione, Traduzione, Commento*, Roma: Edizioni di Storia e Letteratura.

Mazzarino, S. (1966), *Il pensiero Storico Classico*. Vol. 1, Bari: Laterza.

Mignucci, M. (1999), 'Paradoxes', in K. Algra, J. Barnes, J. Mansfeld, and M. Schofield (eds.), *The Cambridge History of Hellenistic Philosophy*, Cambridge: Cambridge University Press, 157-176.

Robinson, T. M. (ed.) (1979), *Contrasting arguments. An Edition of the Dissoi Logoi*, New York: Arno Press.

Stebbing, L. S. (1930), *A Modern Introduction to Logic*, London: Methuen.